Re-examining the Norm Concept in Cephalometrics:

Associations between Subjective Assessments and Deviations from Cephalometric Norms

Joseph de Jesus, D.D.S

Submitted in partial fulfillment of the requirement for the Degree of Master of Science in Dentistry
University of the Pacific, School of Dentistry
San Francisco, California
June 2000

Submitted by ______________________
Joseph de Jesus, D.D.S. M.S.D. Candidate

Accepted by ______________________
Sean Carlson, D.D.S., M.S.

Accepted by ______________________
Roger Boero, D.D.S., M.S.

Accepted by ______________________
Kenneth Snowdowne, Ph. D.
Chairperson, Research Committee

Accepted by ______________________
Robert L Boyd, D.D.S., M.Ed.
Chairperson, Department of Orthodontics

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	22.Aug.00	THESIS

4. TITLE AND SUBTITLE
RE-EXAMINING THE NORM CONCEPT IN CEPHALOMETRICS: ASSOCIATIONS BETWEEN SUBJECTIVE ASSESSMENTS AND DEVIATIONS FROM CEPHALOMETRIC NORMS

5. FUNDING NUMBERS

6. AUTHOR(S)
CAPT DE JESUS JOSEPH Y

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF PACIFIC DENTAL SCHOOL

8. PERFORMING ORGANIZATION REPORT NUMBER
CY00355

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

20000831 097

14. SUBJECT TERMS	15. NUMBER OF PAGES
	48
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank my mentors, Dr. Sheldon Baumrind, Dr. Sean Carlson, and Dr. Roger Boero for their guidance and insights throughout this study. Dr. Carlson's skill in organization, editing, and statistical analysis, as well as his enthusiasm for research, and patience in mentoring were invaluable. Dr. Boero's insights into thorough and coherent writing provided timely guidance throughout the development and presentation of this thesis.

I am also extremely grateful to the six clinicians who volunteered their time to participate in the subjective assessment tasks for this study:

Robert L Boyd, D.D.S., M.Ed.	Donald Poulton, D.D.S.
Ann Marie Gorczyca, D.M.D., M.P.H., M.S.	Richard R. Rutter, D.D.S.
William Lieber, D.D.S.	Vickie Vlaskalic, B.D.S., M.D.S.

Finally, I am indebted to my wife and son, Mary and Danny, for allowing me the time to study, write, and pursue my educational goal.

ABSTRACT

The purpose of this study was to investigate the degree of agreement between deviations from cephalometric norms and clinicians' subjective assessments of treatment difficulty and success. Six orthodontists subjectively assessed and ranked complete pre- and post-treatment records of 36 orthodontic patients for difficulty of treatment and favorability of treatment outcome. Twenty-five hard tissue and 20 soft tissue cephalometric landmarks were located and digitized by two orthodontic residents for all pre- and post- treatment headfilms. Each patient's pre- and post-treatment deviations from 7 commonly used cephalometric hard and soft tissue norm sets were determined. The subjective assessments were compared to the deviations from the published and sample's norms. There were no significant correlations between subjective assessments of difficulty or favorability of treatment and degree of deviations from hard tissue cephalometric norms. Significant negative correlations were found between subjective assessments of favorability of treatment and degree of deviations from specific soft tissue cephalometric norms (E-line and H-Line). This data suggests that clinicians' decisions regarding treatment difficulty or treatment success rely more heavily on factors other than coincidence with hard tissue cephalometric norms.

INTRODUCTION

For over 60 years, norms of hard and soft tissue cephalometric measurements have been used in orthodontics as a primary means of comparing objective information when evaluating patients. In the past, various cephalometric norms have been proposed as "ideal" objective measurements. Many of the commonly used norms, however, are based on limited samples chosen by the personal preferences of a single clinician. Yet even to this day, treating patients to match certain norms is believed to be one way of assuring optimal esthetics and occlusion.

Unfortunately, we do not know how well our traditional cephalometric norms reflect what clinicians subjectively see. Few studies have documented whether there is agreement between clinicians' subjective assessments and deviations from cephalometric norms in estimating treatment difficulty or judging treatment outcome. The primary aim of this study was to examine the association between experienced clinicians' subjective assessments and commonly used cephalometric norms of Tweed, Steiner, Wits, Jarabak, Ricketts, Holdaway, and Arnett.

This study investigated the following:

1. The degree of association between deviations from **hard tissue** cephalometric norms and subjective assessments of treatment difficulty.
2. The degree of association between deviations from **soft tissue** cephalometric norms and subjective assessments of treatment difficulty.
3. The degree of association between deviations from **hard tissue** cephalometric norms and subjective assessments for judging "most satisfactory" and "least satisfactory" post-treatment results.

4. The degree of association between deviations from **soft tissue** cephalometric norms and subjective assessments for judging "most satisfactory and least satisfactory" post-treatment results.

In these areas we proposed the following hypotheses:

1. There is a positive correlation between subjective assessments of treatment difficulty and deviations from **hard tissue** cephalometric norms.

   Reasoning: Traditionally, cephalometric analyses and norms are most refined in the evaluation of anterior-posterior and vertical relationships of dento-alveolar or skeletal components. While many difficulties of treatment do not relate to cephalometric norms (e.g. periodontal defects, missing teeth, arch length deficiencies, etc.), the difficulty of fitting the teeth together within the jaws is commonly assessed via cephalometric norms. We believe those cases judged the most difficult to treat will be the same ones that deviate furthest from the norm.

2. There is a positive correlation between subjective estimates of treatment difficulty and deviations from **soft tissue** cephalometric norms.

   Reasoning: Guiding the changes of soft tissue during orthodontic treatments is unpredictable and difficult. Clinicians normally plan treatment for optimal facial esthetics. Increased deviations from norms may warn a clinician of difficulty in achieving their soft-tissue esthetic goals. It is predicted that clinicians will assess increased difficulty of treatment as the deviations from the norms increases.

3. There is a negative correlation between subjective assessments of favorability of treatment outcomes and deviation from **hard tissue** cephalometric norms.

   Reasoning: If our traditional norms for hard tissues are valid, patients with small deviations should be well treated with outcomes that are judged more satisfactory.

4. There is a negative correlation between subjective assessments of favorability of treatment outcomes and deviations from **soft tissue** cephalometric norms.

   Reasoning: If our traditional norms for soft tissues are valid, patients with small deviations should be well treated with outcomes that are judged more satisfactory.

REVIEW OF LITERATURE

Development of Hard Tissue Cephalometric Norms

The introduction of cephalometric radiography to orthodontics by B. Holly Broadbent in 1931 was a scientific breakthrough in the practice of orthodontics.[1] Prior to the introduction of cephalometric radiographs, orthodontists had no way of objectively measuring the sagittal relationship between the denture, skeleton, and soft tissue profile. Treatment decisions and evaluation of individual treatment results were based almost entirely on subjective philosophies. Cephalometric radiographs promised a source of objective information regarding a patient's skeletal, dental, and facial pattern. Skeletal and dental problems could now be objectively assessed and many clinicians began developing protocols to employ this new tool.

By the 1940s, many clinicians had recognized the importance of gathering objective measurements from cephalometric films. Charles Tweed, in 1946, was one of the first clinicians to publish his views on the link between measurements from cephalometric films and assessments of treatment difficulty and prognosis.[2,3] Tweed's sampling technique was based entirely on personal preference. He sampled individuals he chose as having untreated, "normal occlusions and facial esthetics". He found that certain angles of the mandible and lower incisors, as measured from cephalometric films, tended to have little variation. In his clinical practice, he found that patients who had measurements similar to those of the "normals" tended to be more easily treated to an acceptable occlusion and facial esthetic. Tweed concluded that if a patient's dento-skeletal deviations could be treated to match three specific hard tissue cephalometric measurements (FMPA, FMIA, IMPA), a successful orthodontic result would be assured. Although Tweed's sampling technique was purely subjective and he did not use any type of statistical analysis to support his normative values, his approach of linking a limited number of "normal" cephalometric measurements to treatment planning and estimating esthetic prognosis was a revolutionary trend in orthodontics.

In 1948, William Downs presented one of the first comprehensive methods of hard tissue cephalometric analysis, which utilized a set of normative values.[4] His sampling technique for these norms was purely subjective. He chose a sample of 20 untreated, Caucasian individuals, equally divided as to sex, aged 12-17 with "clinically excellent occlusions". Measurements of their cephalometric radiographs for various angles and linear dimensions were performed and the means calculated. Downs also presented data on the range of variation and standard deviations to these normative numbers, providing a means of determining treatment severity. Two important conclusions of Downs' study were: 1) "There is a facial pattern that represents mean or average form for individuals possessing excellent occlusions." 2) "Single (cephalometric) readings are not important: what counts is the manner in which they all fit together and their correlation with type, function, and esthetics." Unfortunately, Downs did not correlate his normative data with assessments of his sample's facial esthetics.

Arne Bjork, in 1947, presented findings on an extensive cephalometric study of the relationship of craniofacial structures to the facial profile. Utilizing a sample of 322 twelve year-old boys, 281 Swedish army conscripts and a control group of 20 other boys, Bjork developed an extensive list of cephalometric normative values.[5] He also pinpointed the contributions of various skeletal angles to prognathism. His study, however, did not relate cephalometric norms to treatment difficulty, esthetics, or favorability of outcomes.

Cecil Steiner in 1953 presented a cephalometric analysis and a set of normative values that are possibly the most commonly used in orthodontics today.[6] His analysis was synthesized from many of the measurements proposed by previous researchers and concentrated on hard tissue measurements only. Steiner's analysis involved measuring linear distances and angle from specified cephalometric landmarks. He did not state the sample he used to derive his norms, in fact it may have been derived from a single patient. Nor did he relate his norms to assessments of treatment difficulty, esthetics, or post-treatment outcomes. However, the ease of understanding and applying his method to patients has made the Steiner analysis one of the more popular analyses.

Viken Sassouni, in 1955, presented a cephalometric evaluation method based on proportionate analysis of craniofacial structures.[7] The sample on which he based his analysis included 51 girls and 49 boys, aged 7-15. Fifty were Angle class "normal occlusion", 20 class I, 20 class II, and 10 class III. Although Sassouni's analysis involved measurements of particular craniofacial angles and lengths, his concept of the "well-proportioned face" was based on his subjective view of the proper relationships of four horizontal planes and two arcs traced on cephalometric landmarks. No normative data was provided for the sample. In fact, Sassouni warned of the danger of applying any set of norms "indiscriminately to everybody" and concluded that there is no "universal normality."

Alex Jacobson, in 1975, presented a new cephalometric measurement dubbed the "Wits appraisal".[8] It was intended to be a simple diagnostic aid for determining the severity of anteroposterior jaw discrepancies. The Wits measurement reports the positions of the maxilla and mandible relative to the occlusal plane and to each other. Unlike the SNA/SNB/ANB measures, it is independent of variations in cranial base. The norms for the Wits appraisal were derived from a sample of 20 males and 25 females selected on "on the basis of excellence of occlusion". The Wits measurement was not designed to be a full cephalometric analysis. Wits' norms were not related to treatment outcomes or esthetics. However, deviations from the norms were suggested to be indicators of more severe malocclusions.

Development of Soft Tissue Cephalometric Norms

By the late 1950s, clinicians had begun to develop cephalometric analyses to objectively assess soft tissue profiles and relate profile esthetics to underlying hard tissues. In 1958, Charles Burstone described one of the first soft-tissue profile analyses.[9] His method measured fifteen angles of soft tissue profile components from each patient's cephalometric film. Normative data for each of the angles was derived from the Herron sample. The Herron sample consisted of 40 young adult Caucasians with "acceptable faces as chosen by a group of three artists". From his study, Burstone noted that "the soft tissue veneer covering the teeth and bone varies so greatly that study of the dento-skeletal pattern may be inadequate in evaluating facial disharmony". Burstone's "integumental" method of profile assessment was interesting and pinpointed areas of soft tissue deviations, but it did not directly relate these deviations to esthetic outcome assessments or hard tissue measurements.

Robert Ricketts, in 1960, described his recommendations for using cephalometric analysis for clarifying treatment objectives and improving treatment planning.[10,11] He presented a method of cephalometric analysis, that included an objective measurement for relating the lower incisor to profile esthetics. He provided normative data and standard deviations for this objective measurement based on a sample of "1,000 orthodontic cases". He did not note how his sample was chosen. He also proposed a subjective soft-tissue profile line, the "E-line", for improving esthetic treatment planning and assessing treatment results. Ricketts stressed that cephalometric analysis of teeth, skeleton, and the overlying soft-tissues are required to set proper treatment objectives for "ultimate functional and esthetic balance and harmony of the teeth, mouth, jaws and face". Ricketts' cephalometric analysis related one hard tissue (lower incisor) and one soft-tissue (E-line) measurement to profile esthetics but did not relate other cephalometric norms to treatment difficulty or assessment of esthetic treatment results.

To provide a "more specific guide to help young orthodontists achieve the maximum facial harmony" in their orthodontic treatments, L. Levern Merrifield, introduced a method of relating hard and soft-tissues in a cephalometric analysis.[12] In 1966, Merrifield devised the "Z-angle" or "profile line" and referenced it to the "Tweed triangle" of hard tissue measurements. He presented the Z-angle as a helpful cephalometric angle for objectively evaluating facial esthetics. His normative data was based on a sample of 40 "Tweed non-orthodontic normals", 40 "orthodontic (treated) normals" from Tweed's practice, and 40 "orthodontic normals" from his own practice. He selected a sub-sample of the ten best faces from each group, based on his criteria of "soft-tissue contour, balance and pleasing esthetics". Merrifield provided normative data for the "Z-angle" for this group and explained that this measurement could provide a "critical description…and eliminate the vagueness of 'eye judgment'".

One of the more comprehensive soft-tissue cephalometric analyses with 11 measurements and corresponding normative data was introduced by Reed Holdaway in 1983.[13,14] The sample he used to derive his technique and norms is unclear, as it came from "years of observation and description of patients from the private practice of the author". It is probable that he derived his norms from attractive profiles selected by personal preference. Holdaway described the use of his analysis and norms in orthodontic treatment planning by outlining a step-by-step "visual treatment objective" method which is used, in various forms, by many clinicians today. Several of his soft-tissue cephalometric measurements and norms were related directly to underlying hard tissue measurements. However, he did not correlate his soft-tissue measurements with cephalometric norms or assessments of facial esthetics or difficulty of treatment.

In a 1993 article, G. William Arnett, an oral-maxillofacial surgeon, remarked on the problems of relying on cephalometric hard tissue analyses when diagnosing and planning facial esthetics. He noted that "correction of the bite does not always lead to correction, or even maintenance, of facial esthetics" because of the great variability of soft tissues.[15,16] He presented a comprehensive clinical exam technique for analyzing facial esthetics to avoid potential facial imbalance from orthodontic/surgical treatments. Following on this work, in 1998, Arnett presented a new 46 measurement, soft tissue cephalometric analysis (STCA) based on natural head position and a "True Vertical Line" to subnasale. Arnett's measurements and norms were derived from 46 adult white models (20 male, 26 female) with Class I occlusions viewed by Arnett as "reasonably facially balanced".[17] Arnett's analysis is comprehensive in relating soft-tissue position and thickness with the spatial position of underlying dentoskeletal components. His study did not relate his norms to other clinicians' subjective views of facial esthetics. Nevertheless, Arnett's soft tissue cephalometric analysis and norms for evaluation of facial balance and harmony are considered by many to be "state of the art" for orthognathic treatment planning and outcome assessment.

Testing the Relationship between Norms and Esthetics

At the same time orthodontists were striving to perfect methods of cephalometric analyses for hard and soft tissues, others were testing the relationship between cephalometric measurements and assessments of esthetics. Richard Riedel, in 1950, was probably one of the first to study and publish the relationship between orthodontists' opinions on profile esthetics and cephalometric measurements.[18] Using profile tracings from the cephalometric films of 29 treated and untreated patients, 69 orthodontists were found to be "remarkably uniform" in their opinions when judging "good" or "poor" profiles. 40 cephalometric measurements were then applied to the 13 patients judged to have "good" and "poor" profiles. Riedel found that "common differences" in eight of the skeletal and dental measurements existed between profiles judged "good" and "poor". He did not provide normative values for these measurements, but he did conclude that the proper relation of these dento-skeletal components were "important in esthetic balance". Unfortunately, measurements of correlations and statistical significance between the collected opinions and cephalometric measurements were not performed in his study.

Wendell Wylie , in 1955, compared the relationship of profile esthetics to maxillary and mandibular central incisor position.[19] Wylie noted anecdotally that most orthodontists when selecting faces they like chose those with straighter profiles. Therefore, he limited his study by defining improved profile esthetics in this group as a tendency toward straightened profiles. Using a sample of 29 patients consecutively treated by Tweed, he measured the change in mandibular incisor to Frankfort-horizontal and change in maxillary incisor to S-N and correlated these with change in the soft tissue angle of convexity. Wylie found very low correlation between changes in FMIA and increases in esthetic straightening of the profile. He also found low correlation between changes in the maxillary incisor position and profile straightening. . He concluded there is "no clear cut relationship between soft tissue profile changes and inclination of the incisors". Wylie's study was limited in scope to an investigation of a cephalometric "ideal" of the Tweed Triangle and it did not objectively address other cephalometric norms or clinicians' opinions on esthetics.

Richard Reidel, in 1957, once again studied the relationship between facial esthetics and cephalometric measurements. Additionally, he compared a sample's measurements to established cephalometric norms. Reidel selected a sample of 30 "Seattle Seafair princesses and their queen".[20] These individuals were chosen because their facial appearance was deemed to be a "more or less ideal young female type" by non-orthodontist judges. Reidel measured the sample's head films for cephalometric angles of Downs, Tweed, and Steiner, then calculated means and ranges for his sample and compared them to "established" norms. Two important conclusions were noted: 1) "The skeletal patterns of the girls studied were very similar to those recorded in previous studies of normal occlusions." 2) "The public's concepts of acceptable facial esthetics are apparently in good agreement with standards established by orthodontists on the basis of normal occlusion." This study supported the idea that facial esthetics and cephalometric norms are linked, however statistical correlations were not performed.

The statistical relationship between angular cephalometric measurements and facial appearances of untreated orthodontic patients was reported by Donald Poulton in 1957.[21] A sample of 28 boys and 37 girls consecutively accepted for orthodontic treatment was used. Ten University faculty orthodontists subjectively ranked, from best liked to least liked, lateral facial photographs of the girls, then the boys. The highest and lowest eight in ranking for each group were used for the statistical comparison. Cephalometric measurements were made for the 24 ranked individuals. Ten angles from analyses of Downs, Steiner, Tweed, Riedel and Grusd were measured. Correlation via a Wilcoxon test showed that certain angular measurements "differentiated between the extremes in facial esthetics as judged here". Poulton's study suggested that 4 specific cephalometric angles and norms, because of their correlation with facial pattern, would be more useful in headfilm analysis.

Harvey Peck, in 1970, investigated the general public's concept of facial esthetics in relation to the cephalometric measurements commonly used by orthodontists.[22] A sample of 49 female and three male Caucasian young adults with "publicly recognized American faces" (i.e. models, beauty contest winners and actresses/actors) were chosen. The sample was cephalometrically analyzed using the Margolis, Downs and Steiner analyses and mean data calculated. Peck found that "most of the 29 measurements fell within range of the pre-established standards" (or norms), however the sample means "favored a fuller, more protrusive dento-facial pattern than our cephalometric standards would have liked to permit." Statistical correlations were not given. Peck also described a "profilometric" analysis for evaluating facial esthetics from patients' photographs and gave mean and deviation values for various soft-tissue angles measured from the same sample group. He recommended this analysis for objectively assessing facial esthetics rather than relying on cephalometric analyses and norms.

Young-Chei Park and Charles Burstone, in 1986, presented a study that tested the "efficacy" of using a cephalometric norm as a basis in orthodontic treatments for producing "predictable and desirable facial esthetics".[23] They chose a random sample of 30 adolescents from the practices of orthodontists who primarily used hard-tissue criteria for treatment planning. The patients selected were those who at the end of successful treatment had lower incisors 1.5 mm anterior to the A-pogonion plane. They compared extensive cephalometric hard and soft tissue measurements of the sample group with the same measurements for an esthetically pleasing control group of 32 adolescents selected by the authors "on the basis of good facial form". Several measurements and overall profile esthetics (as determined by accepted soft tissue cephalometric norms) differed greatly between the two groups. In particular, Park and Burstone found that the sample group showed a large variation in the position of the lips when treatment to the cephalometric dento-skeletal standard was performed. They equated this large variation with inconsistent esthetic results. They concluded, "any given dento-skeletal standard has questionable validity in producing either desirable esthetics or reproducible profiles following treatment."

In his Masters thesis, Jack DuClos, in 1995, studied the relationship between borderline patients' cephalometric measurements and orthodontists' subjective decisions regarding extraction versus non-extraction treatments.[24] He also studied the relationship between his sample's post treatment cephalometric measurements and the clinicians' assessments of facial esthetics and improvement from treatment. His study did not investigate the relationship between deviations from cephalometric norms and clinicians' assessments of pre treatment difficulty. His sample included 48 adolescents between age 10-15 at the start of treatment, chosen at random from the private practice of an experienced orthodontist. His study used selected cephalometric measurements of Downs, Tweed, and Steiner. Fifteen experienced orthodontists provided subjective assessments of the pre- and post-treatment records for the patients. DuClos found that their assessments of post-treatment facial esthetics showed significant association with soft-tissue facial convexity measurements, but not hard tissue convexity measurements. He noted that soft-tissue variations could mask underlying skeletal discrepancies, which might otherwise detract from facial esthetics. DuClos also found that the subjective assessments of improvement from treatment were significantly associated with decreases in hard and soft tissue facial convexity (i.e. straighter profiles) and decreases in vertical skeletal relationships. He suggested that profile facial appearance is especially linked to relative harmony of nose and chin and balance of vertical facial thirds and restated the importance of analyzing both hard and soft tissues in treatment planning.

MATERIALS AND METHODS

Sample

Our sample was originally collected as part of another study completed in 1995 by Dr. David Teeter.[25] In his study, all cases started at the University of the Pacific from 1988 through 1992 were reviewed. Inclusion in the sample was based on two criteria:

1. Cases had to have complete pre- (T1) and end-treatment (T2) records, including facial and intra-oral photographs, study casts, lateral cephalometric films, and full-mouth intra-oral or panoramic radiographs.
2. The case had to be treatment planned under the supervision of one of the six orthodontic faculty participating in Dr. Teeter's study. These faculty were all active clinical instructors at UOP from 1988 to 1995.

Of the 630 cases reviewed, 162 met the criteria. Each pre-treatment record set was judged by at least 2 faculty participants as either extraction or non-extraction treatment cases. Inter-clinician disagreement cases, 44 total, were noted to be "borderline" extraction/non-extraction cases. The borderline status of these cases was incidental in the selection of these cases for the current study. This subset of cases was used primarily because the cases had already been vigorously screened for inclusion in the Teeter study (i.e. patients treated without orthognathic surgery, presence of complete pre- and end-treatment records), extensive cephalometric landmark data required for the study was already acquired, and the number of subjects provided proper statistical power for the desired correlation study.

For our study, 36 of the 44 borderline cases from the Teeter sample were randomly selected. A goal of the sampling method was to select 18 extraction and 18 non-extraction cases. Random numbers were assigned to the entire 44 cases. These numbers were then randomly selected until 18 extraction and 18 non-extraction cases were selected.

The sample of 36 was again divided into 3 groups of 12 cases to make the assessment/ranking task manageable for our judges. We attempted to have equal numbers of patients having received extraction and non-extraction treatments in each group. The first group of 12 cases however, had 7 extraction and 5 non-extraction treatments, because one of the non-extraction cases originally assigned did not have adequate records at the time of subjective data acquisition.

Simple demographics for our sample are shown in **Table 1**. Demographic information for each case is provided in *Appendix A*.

Sample Size	Males	Females	Extractions	Non-Extractions	Average age at T1	SD	Average age at T2	SD	Average Treatment Time	SD
36	16	20	17	19	17.30	8.42	20.13	8.39	2.83	0.99

Table 1: Sample demographics

Cephalometric data acquisition

Cephalometric landmarks for each of the 36 cases for both T1 and T2 films were independently determined by two University-trained, graduate orthodontic residents (N.N. and J.R.). Several steps were taken to calibrate the graduate students prior to their identification of the sample landmarks. The students were given a review of the anatomy, definitions, and method of landmark location. Two calibration sessions were then completed in which each student independently identified the 25 landmarks on 3 cephalometric films. Following the first session, differences in landmark definitions and location technique were noted and clarified. The second session was used to verify the calibration between the two students. The procedure for cephalometric head film analysis and digitization followed the protocol described by Baumrind, et al.[26] Fifteen skeletal landmarks and 10 dental landmarks were identified for all 36 cases (**Figures 1-2**). Twenty soft-tissue landmarks for the 36 cases were also identified and digitized in a previous study by Nalchajian, et al.[27] (**Figure 3**) Fiducials and landmarks were digitized into the computer in the form of Cartesian coordinates referenced to the Frankfort Horizontal plane. The software program ELLIPSE was used to compare the location of each tracer's landmarks in relation to the envelope of error for each given landmark. A previous study by Baumrind,et al. described the characteristic envelopes of error that correspond to location of landmarks.[28] When the two tracers were both within the envelope of error for a given landmark, the computer recorded the landmark by averaging the X and Y coordinates (AVEPIC). When one or both of the tracers were outside the envelope of error, the given landmark was re-traced. If the landmarks were outside the envelope a second time, the landmark was assumed to be indistinguishable due to poor film quality. Independent estimates were averaged and the average coordinate values were loaded into a SQL server database. Cephalometric angles and linear measurements were computed from the averaged landmark values using specialized computer software.

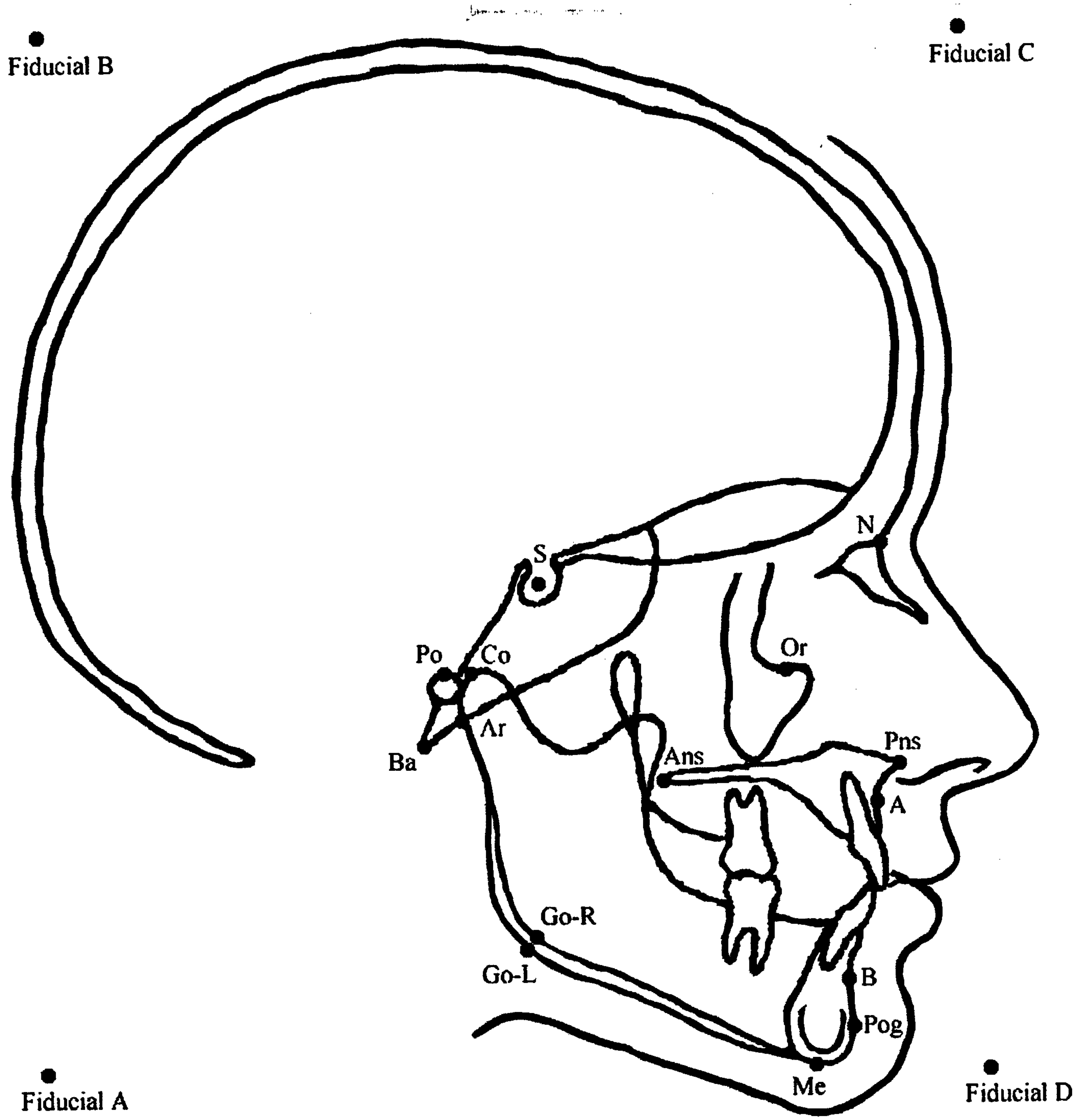


Figure 1: Digitized cephalometric hard tissue landmarks

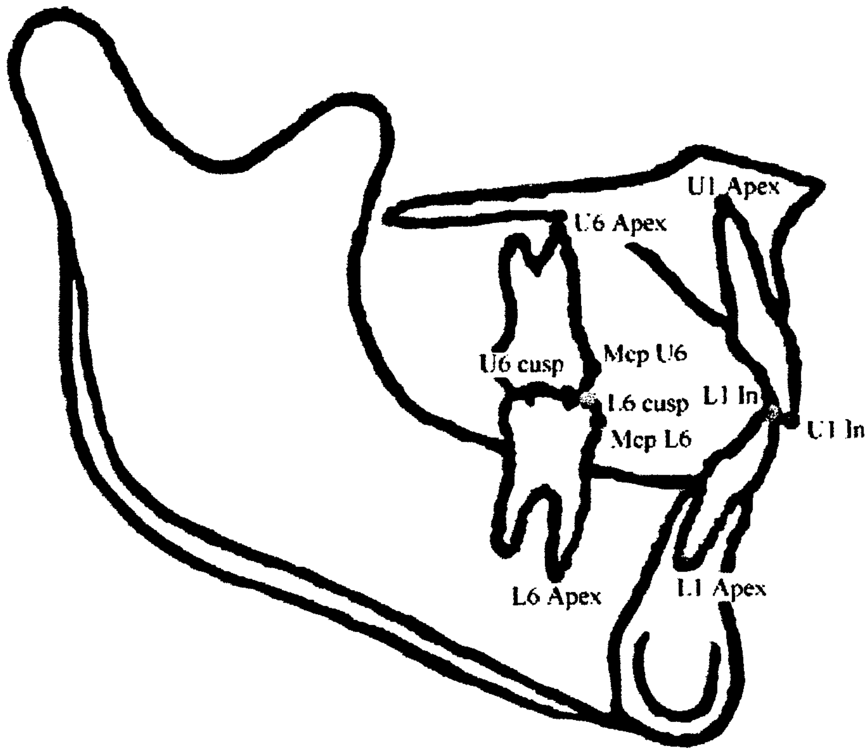


Figure 2: Digitized cephalometric dental landmarks

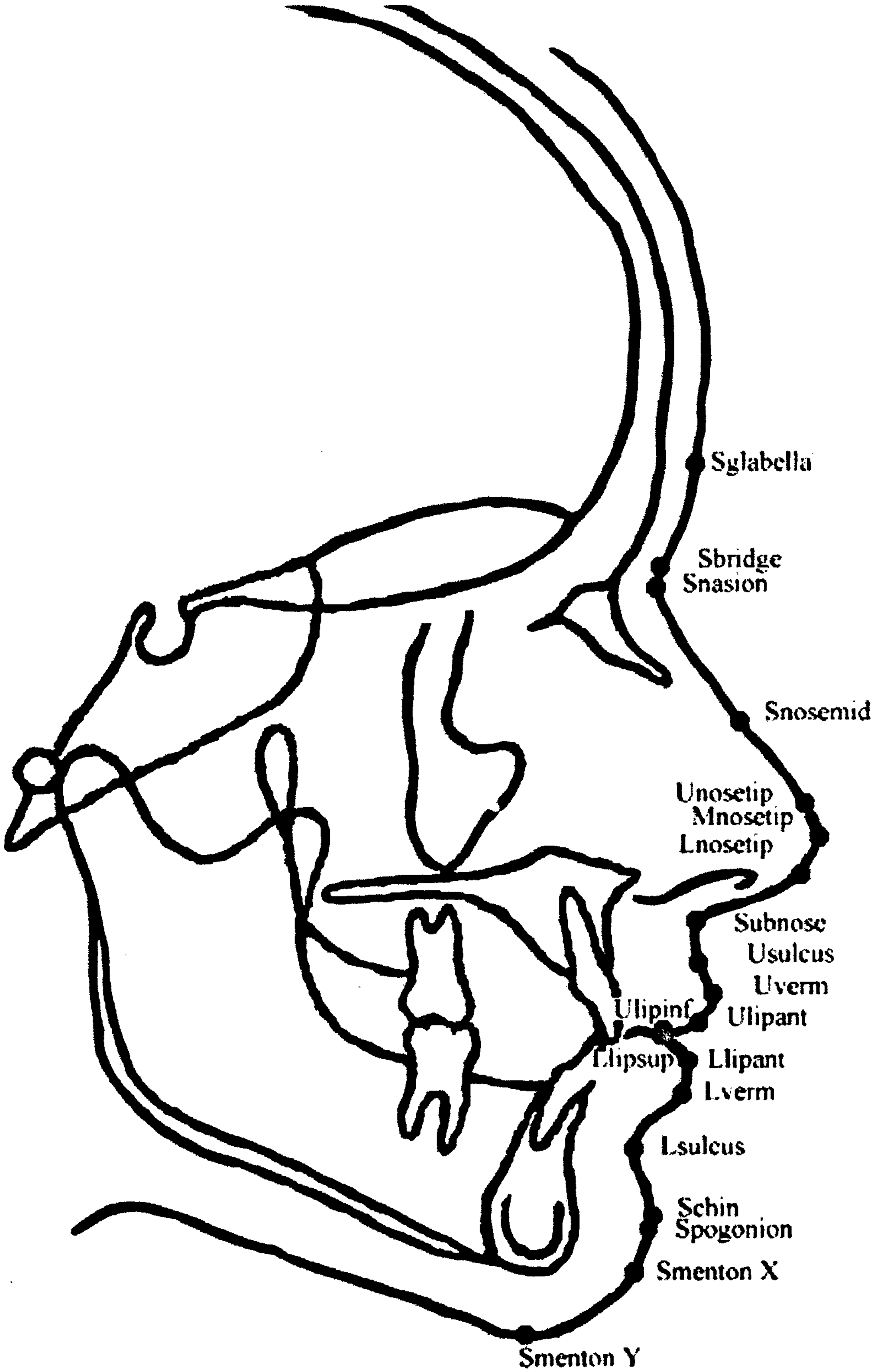


Figure 3: Digitized cephalometric soft tissue landmarks

Deviations from published cephalometric norms for the following analyses **(Figures 4-10)** were investigated:

Hard Tissue:

1. Tweed
2. Steiner
3. Wits
4. Jarabak
5. Ricketts

Soft tissue:

1. Ricketts
2. Holdaway
3. Arnett

Cephalometric norms for certain analyses were originally published with a range, rather than a distinct number. The norm used in these cases was the median number in the normative range.

Subjective assessment data acquisition

Four faculty clinicians were recruited based on several criteria: Full-time, tenured faculty members were recruited based on years of clinical and academic experience and projected availability in the future for similar comparison studies. Two additional volunteer faculty clinicians were recruited to provide a wider range of experience and educational background. **Table 2** provides demographic information on the 6 judges.

Name	Grad date	Ortho school	Yrs private practice	Yrs teaching	ABO cert.
Boyd	1974	U of Penn	25	23	Yes
Gorczyca	1990	Northwest	9	9	No
Lieber	1974	UOP	25	23	Yes
Poulton	1956	UCSF	40	40	Yes
Rutter	1961	Columbia	35	40	Yes
Vlaskalic	1994	U of Melbourne	2	3	No
AVG			22.7	23	

Table 2: Demographics of Judges

The six judges were separately presented with each of the 3 sets of twelve T1 case groups, one group at a time. The records were displayed identically for each case (**Figure 11-12**). Names were masked on all records. The judges were asked to review all of the records for each case. Subjective assessment data was collected from each of the judges in the form of evaluation sheets (**Figure 13-14**) for the following 2 tasks:

Task 1: Evaluation of full pre-treatment T1 records (lateral head film, FMX, initial photos, and study models) for initial clinical impression of estimated difficulty, time of treatment, likelihood of extraction, and surgical likelihood. Head film tracings were not included in the record set. Independently, each judge was required to rank the T1 difficulty of each group of 12 cases into three equal subjective groupings of 4 "least difficult", 4 moderately difficult, and 4 "most difficult" cases.

Task 1 was completed for the three experimental groups before initiating Task 2. This allowed several weeks to elapse between assessment of the initial T1 and end-treatment T2 records. Similarly as in Task 1, the six judges were separately presented with each of the 3 sets of 12 T2 case groupings, on group at a time. Groupings of Task 2 cases were unchanged from Task 1.

Task 2: Evaluation of post-treatment T2 records (lateral head film, end-treatment photos, and study models) for "most satisfactory" and "least satisfactory" treatment outcome. Head film tracings were not included in the record set. Independently, each judge was required to rank the T2 favorability of each grouping of 12 cases into 4 "least satisfactory", 4 satisfactory, and 4 "most satisfactory" cases.

Subjective assessment scores gathered using the evaluation sheets were entered into an Excel spreadsheet to archive them and the data was prepared for the SAS program. Cases ranked "most difficult" at T1 and "most satisfactory" at T2 were assigned a rank score of "3" in the Excel spreadsheet. Scores of "2" correspond to rankings of "moderately difficult" at T1 and "satisfactory" at T2. Scores of "1" indicate rankings of "least difficult" at T1 and "least satisfactory" at T2. Rank scores from all six judges were added for each case to calculate the total assessment score for each case at T1 (difficulty) and T2 (favorability). A Total Assessment Score (TAS) of "18" means six judges each ranked a case as "3" or unanimously most difficult at T1 or most satisfactory at T2.

Statistical Analysis

SAS software package was used to calculate the "r" value for correlation of the cephalometric deviations with 1) published norms and 2) our sample's norms with the Total Assessment Score for each of the 36 cases. Correlation (r) and confidence level (p) were determined. Correlations of <0.25 indicate little or no relationship, 0.25-0.50 fair degree of relationship, 0.50-0.75 moderate to good, and >0.75 very good to excellent. [29] Confidence level for our study was set at p=0.05.

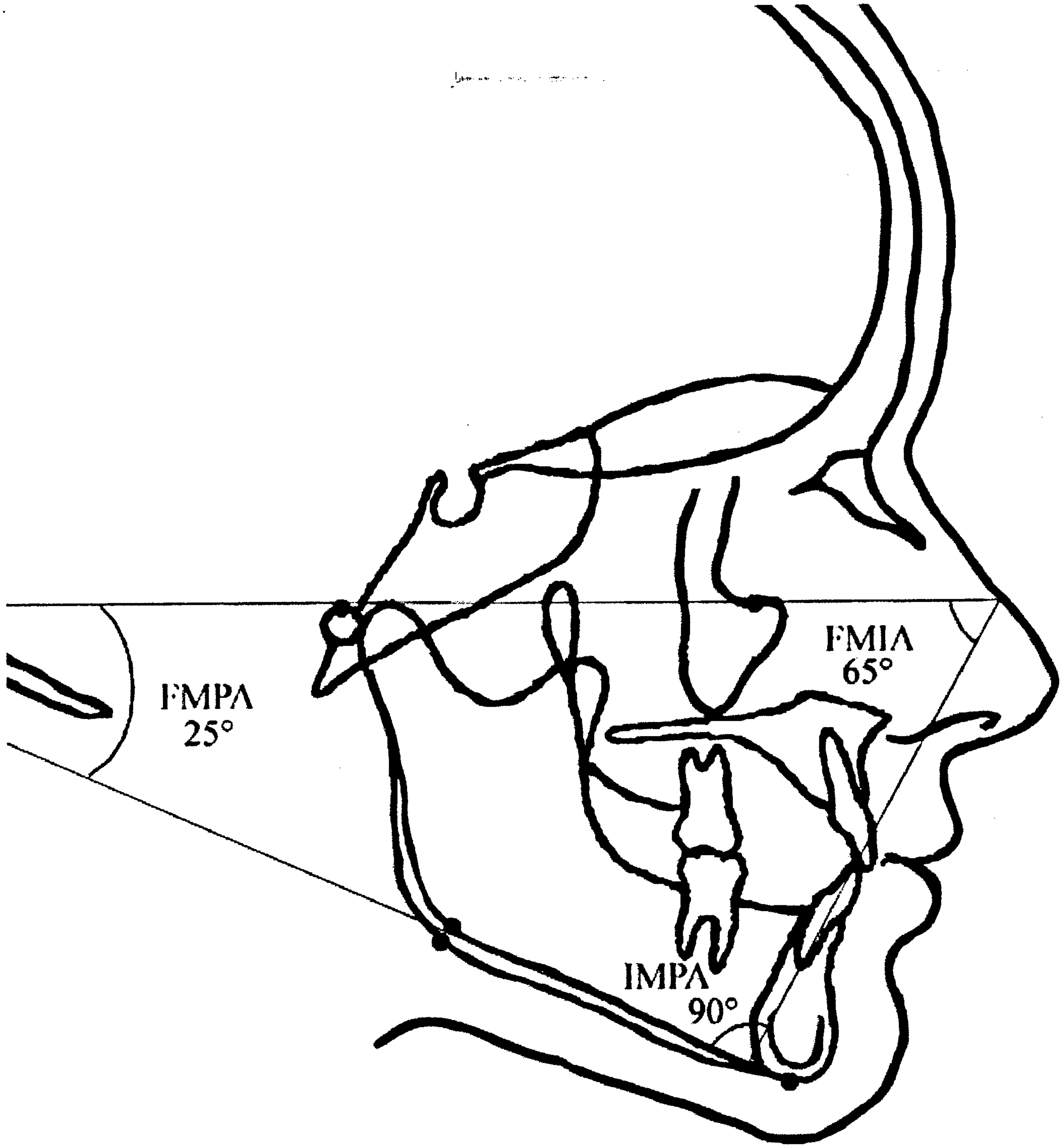


Figure 4: Tweed cephalometric measurements

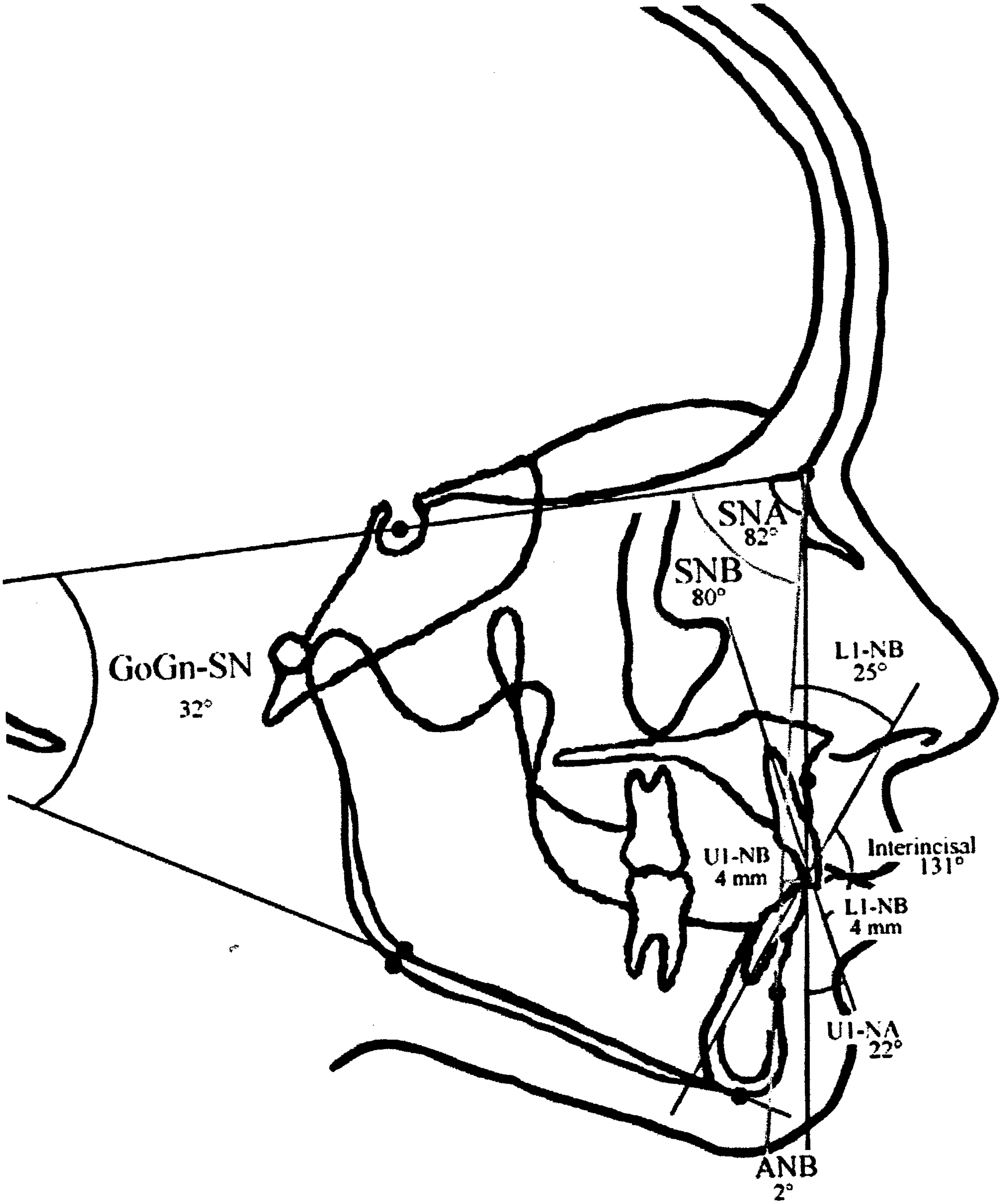


Figure 5: Steiner cephalometric measurements

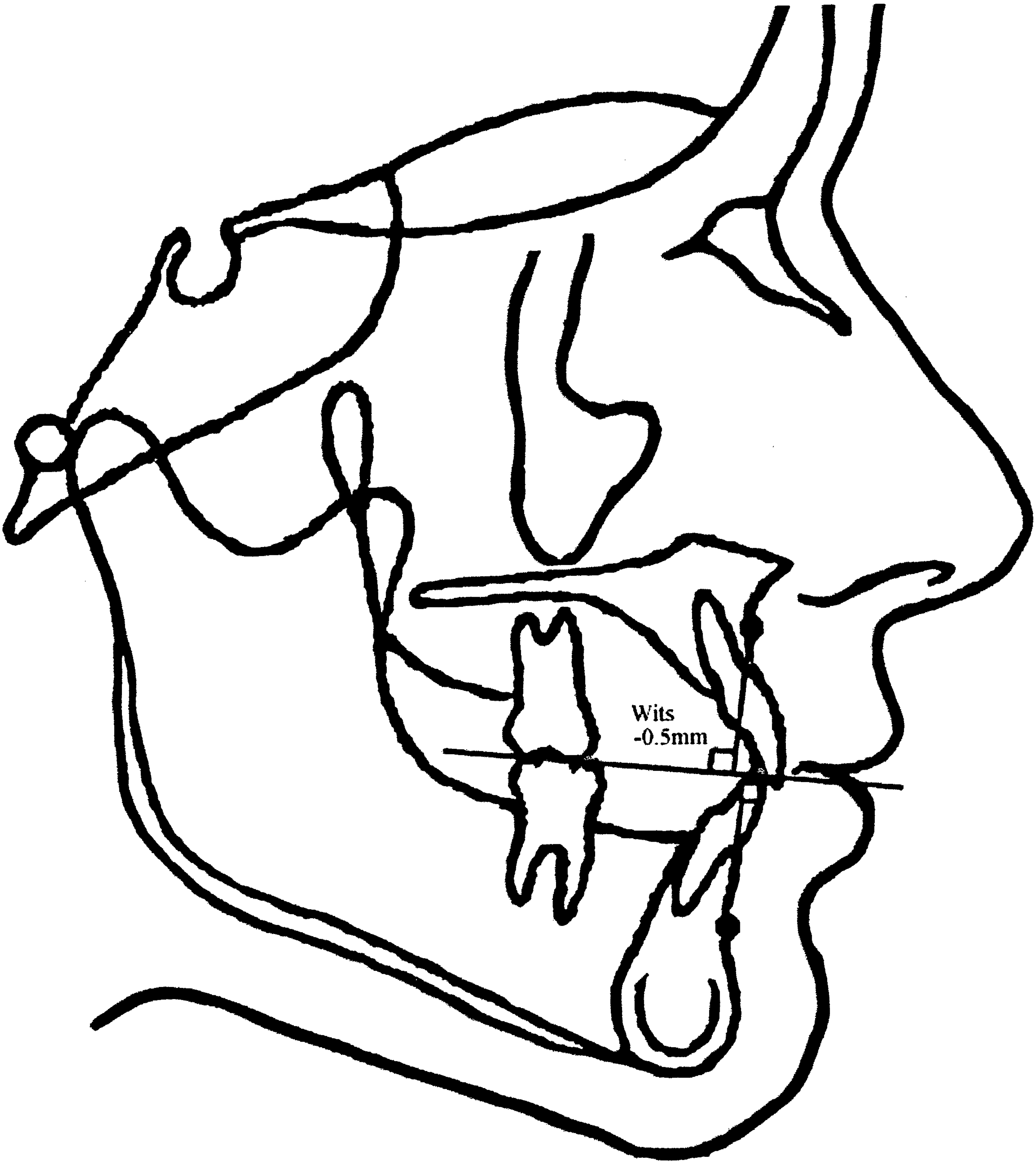


Figure 6: Wits cephalometric measurement

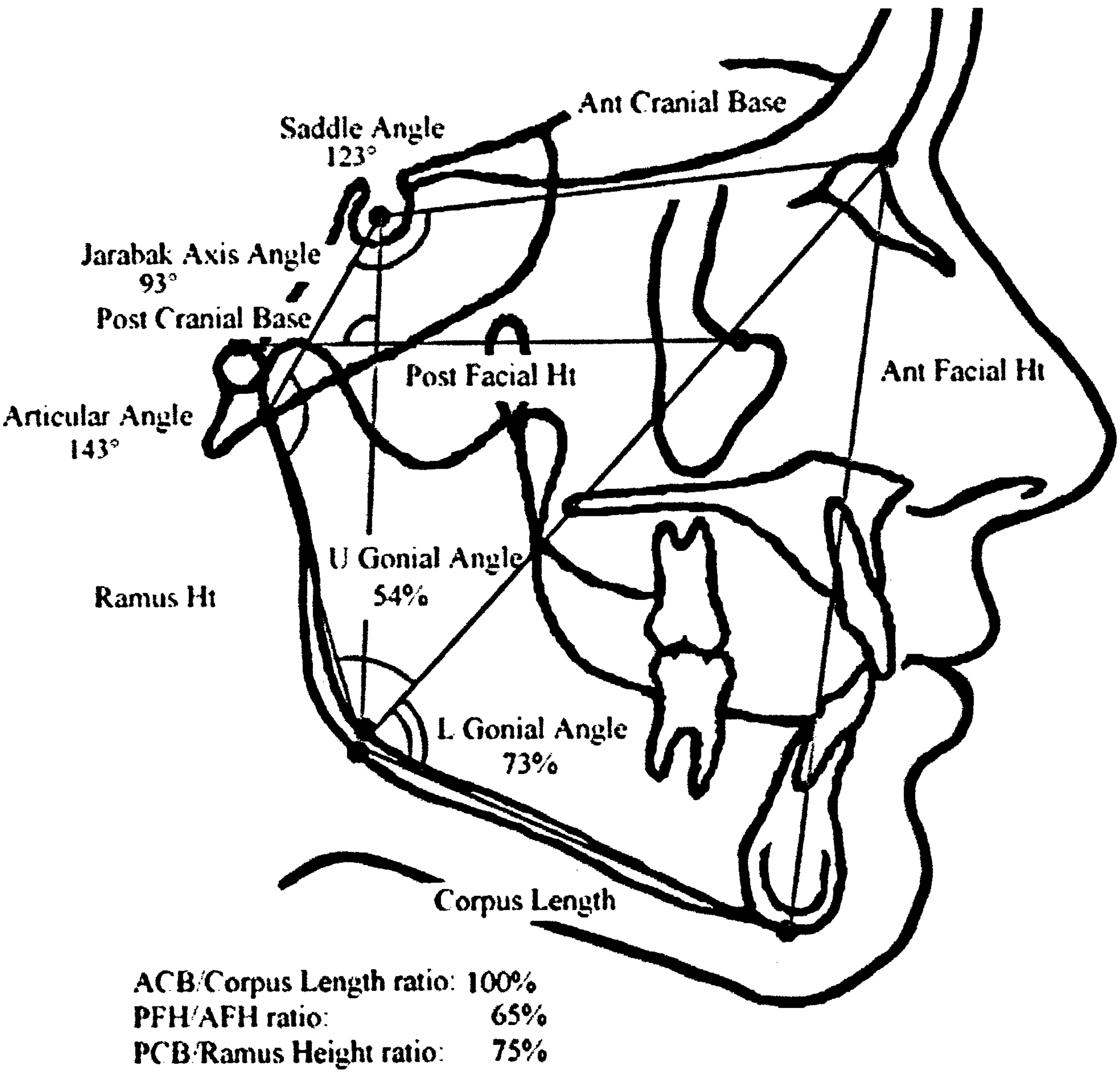


Figure 7: Jarabak cephalometric measurements

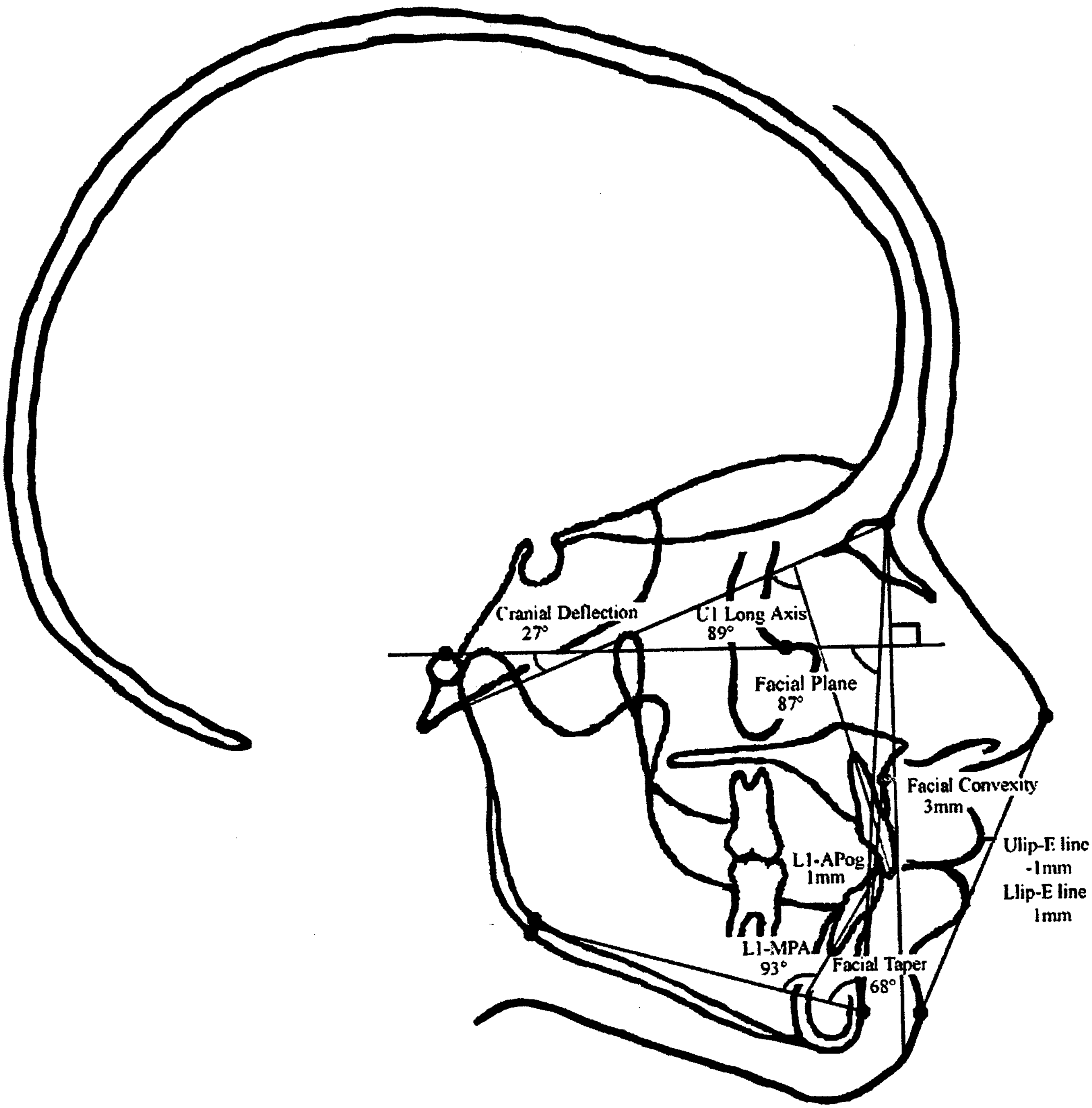


Figure 8: Ricketts cephalometric measurements

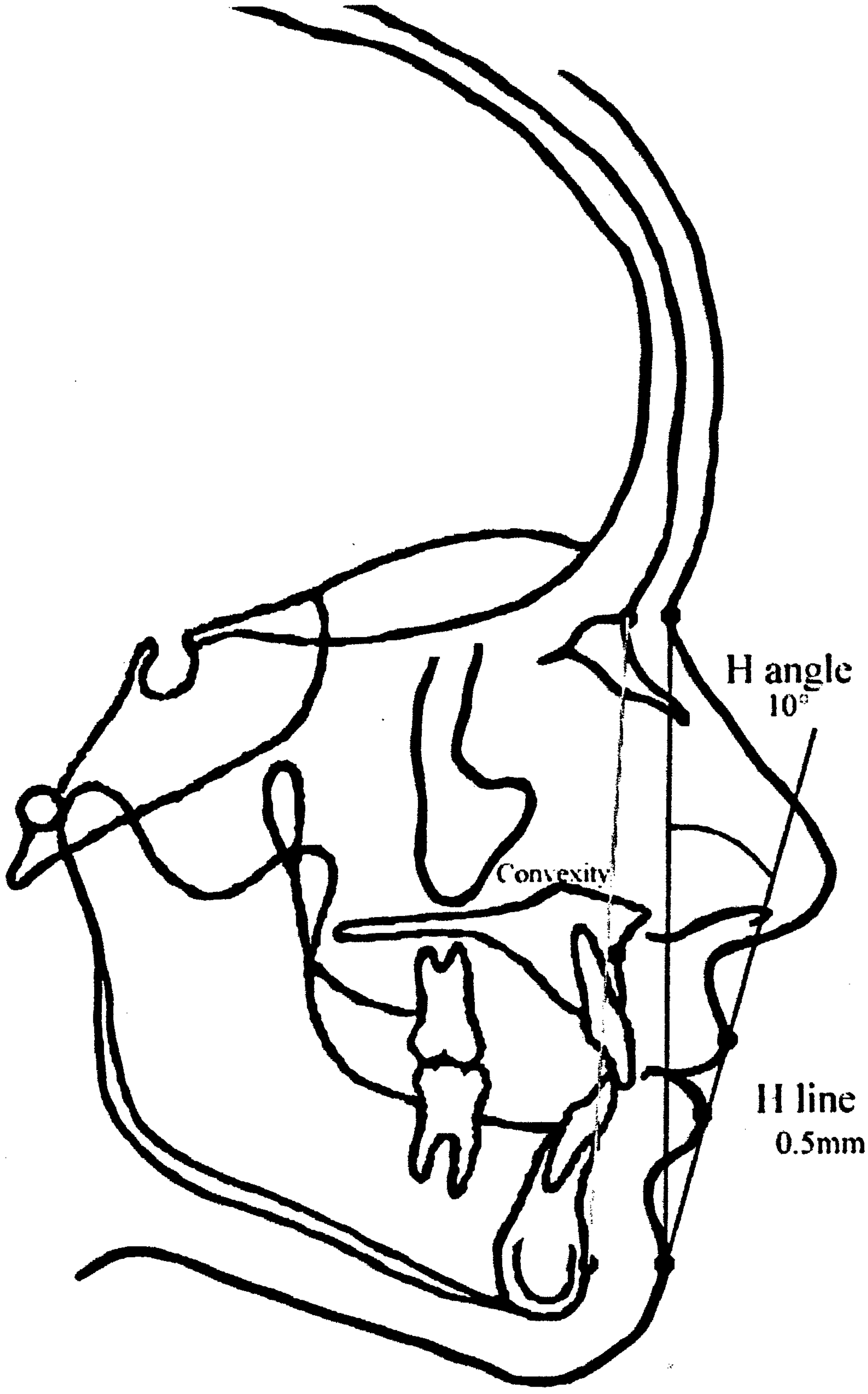


Figure 9: Holdaway soft tissue cephalometric measurements

harmony

	F	M	
total face (facial angle)	169.3	169.4	total facial harmony
forehead to Mx (G'- A')	8.4	7.8	
forehead to chin (G'-Pog')	5.9	4.6	
Sn to chin (Sn-Pog')	3.2	4.0	interjaw harmony
Mx to Md (A'-B')	5.2	6.8	
lip to lip (ULA-LLA)	1.8	2.3	
lower incisor to chin (Md1-Pog')	9.8	11.9	intra mandibular harmony
lower lip to chin (LLA-Pog')	4.5	4.4	
Md to chin (B'-Pog')	2.7	3.6	

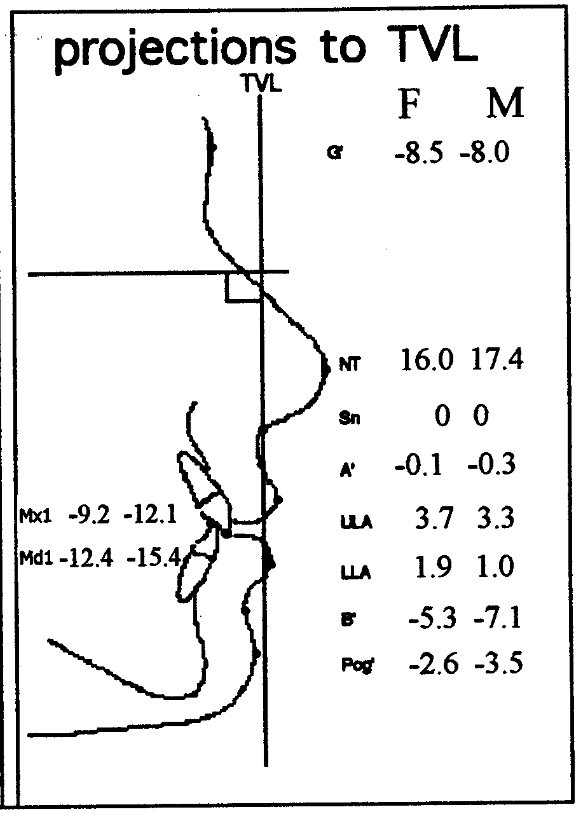
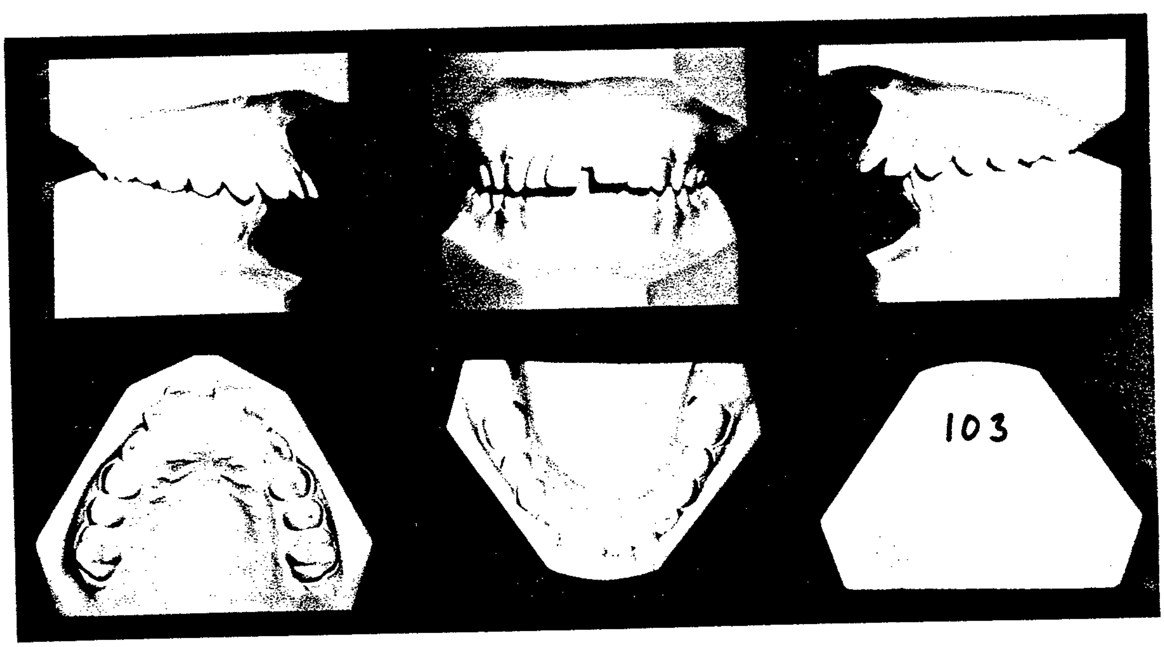


Figure 10: Arnett Soft Tissue Cephalometric Analysis

Figure 11: Sample T1 records displayed to judges. Original casts, headfilms, and panoramic or full mouth radiographs were presented.



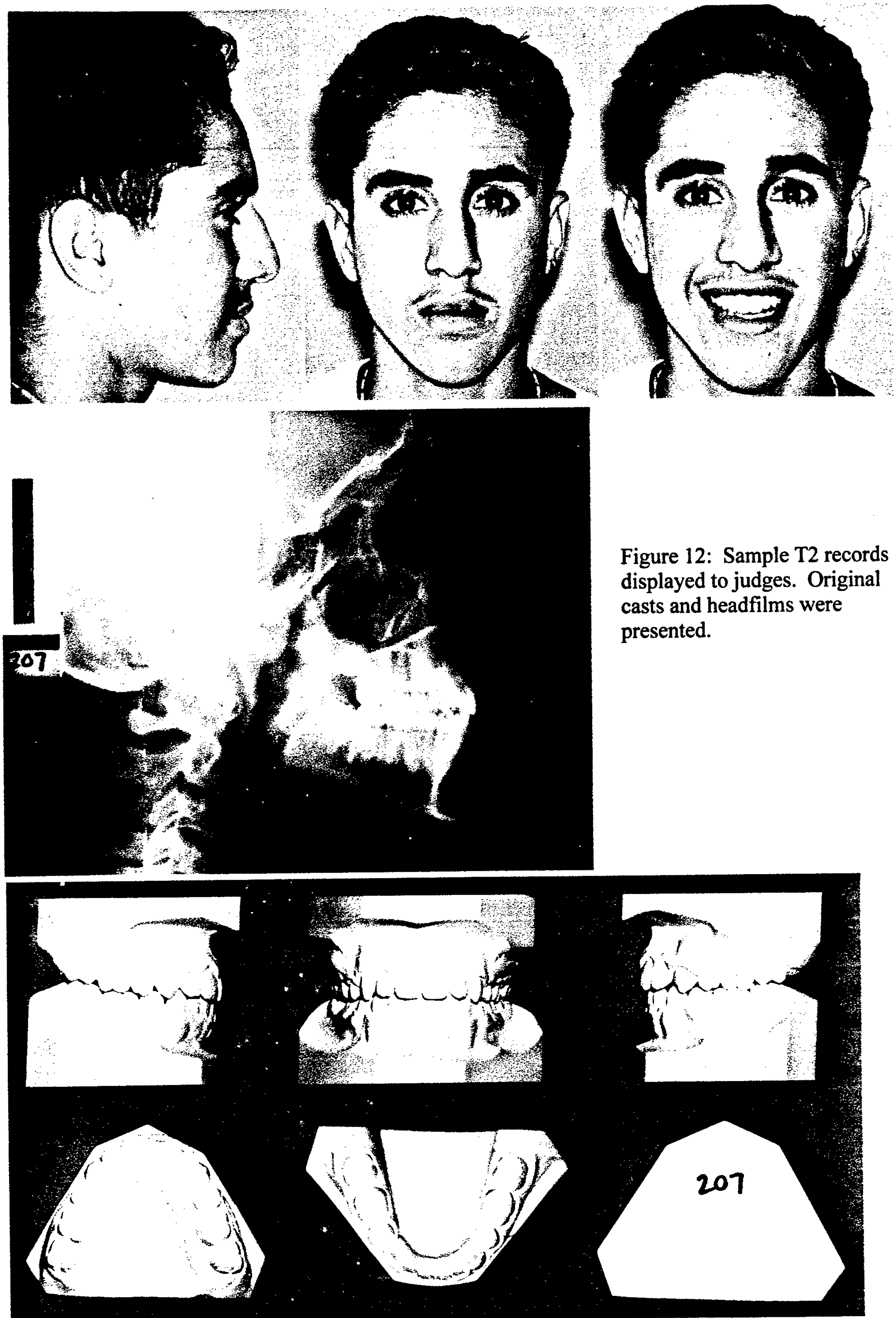


Figure 12: Sample T2 records displayed to judges. Original casts and headfilms were presented.

de Jesus Task 1

Date: ____________ Investigator Name: ____________ Sample Group: #1 (of 3)

Craniofacial Research Instrumentation Laboratory
UOP-Department of Orthodontics-Clinical Decision Making Study
INITIAL CLINICAL IMPRESSION

Pre-treatment cephs, panos/FMXs, study casts, and facial photographs are arrayed on this table for twelve subjects.

1. Please review these records and fill out the table below for each patient.

2. After you have finished, please select the four most difficult and the four least difficult cases.

Case I.D. #	Angle Classification? (Cl I. II. III+ division and subdivision)	X or NX treatment? (Please indicate X or NX)	If X, which teeth would you extract? (4|4 = upper bicuspids)	What is your estimate of treatment time in months?	Is orthognathic surgery indicated? (x for each case where indicated)	Four most difficult (x only four)	Four least difficult (x only four)
101							
102							
103							
104							
105							
106							
107							
108							
109							
110							
111							
112							

Figure 13: Task 1 Subjective Assessment data sheet for T1 case groupings

Task 2

Date: ____________ Investigator Name: ____________ Sample Group: #1(of 3)

Craniofacial Research Instrumentation Laboratory
UOP-Department of Orthodontics-Clinical Decision Making Study
EVALUATION OF TREATMENT OUTCOMES

Post-Treatment cephs, study casts, and facial photographs are arrayed on this table for twelve subjects who received full bonded orthodontic treatment.

Please examine the records for all twelve subjects and record your choice of the **four** "Most Satisfactory" and the **four** "Least Satisfactory" in the spaces provided below. Please indicate briefly no more than **three reasons** for each choice.

A. MOST SATISFACTORY

Case I.D. #____	Case I.D. #____	Case I.D. #____	Case I.D. #____
1.________	1.________	1.________	1.________
2.________	2.________	2.________	2.________
3.________	3.________	3.________	3.________

B. LEAST SATISFACTORY

Case I.D. #____	Case I.D. #____	Case I.D. #____	Case I.D. #____
1.________	1.________	1.________	1.________
2.________	2.________	2.________	2.________
3.________	3.________	3.________	3.________

Figure 14: Task 2 Subjective Assessment data sheet for T2 case groupings

RESULTS

Table 3 shows the sample's cephalometric means and ranges for pre-treatment T1 and end-treatment T2 headfilms. The published norms for the analyses used are also present for comparison.

Subjective data on initial T1 difficulty and end-treatment T2 favorability of each case is reported in **Table 4** as Total Assessment Score (TAS). A Total Assessment Score of 18 indicates greatest relative difficulty at T1 and greatest relative favorability at T2 as assessed by the judges. Note that 2 cases were judged unanimously most difficult and 5 cases were judged unanimously most favorable. *Appendix B* provides raw data on the individual judges scores of each case for T1 and T2, as well as reasons stated for assessment of greater T2 favorability for various cases.

Table 5 shows the correlation (r-value) and probability values (p) when comparing the sample's individual case deviations from the published norms with the Total Assessment Score for individual case difficulty and favorability. Raw data from the SAS program for these correlations and probabilities are provided in *Appendix C.*

No significant correlations were demonstrated between the individual case TAS for difficulty and degree of case deviation from the published hard tissue norms of Tweed, Steiner, Wits, Jarabak, and Ricketts. **No significant correlations** were shown between the individual case TAS for difficulty and degree case of deviation from the published soft tissue norms of Holdaway, Ricketts, or Arnett.

No significant correlations were noted between the individual case TAS for favorability and degree of case deviation from published hard tissue norms for Tweed, Steiner, Wits, Jarabak, and Ricketts. **Significant negative correlations** were discovered between the individual case TAS for favorability and degree of case deviation from the published soft tissue norms for

Ricketts' E-line (lower lip) and Holdaway's H-line, i.e. decreased deviation from the published norms was associated with increased favorability. No significant correlations were noted for deviations from Arnetts' soft tissue norms.

Table 6 indicates the correlation and probability statistics when comparing sample's individual deviations from the total sample mean (or norm) with the Total Assessment Score for individual case difficulty and favorability. Raw data from the SAS program for these correlations and probabilities are provided in *Appendix D.*

No significant correlations of individual case TAS for difficulty with degree of deviation from the hard tissue sample norm were shown when using the analyses of Steiner, Wits, Jarabak, and Ricketts. A **significant negative correlation** was shown between individual case TAS for difficulty and degree of deviation of the Tweed FMPA measurement from the sample norm. Surprisingly, this shows that increased deviation from the sample's norms was associated with decreased difficulty! Individual case TAS showed **no significant correlation** with degree of deviation from soft tissue sample norms using the analyses of Ricketts, Holdaway, and Arnett.

No significant correlations were demonstrated when comparing individual case TAS for favorability with degree of deviation from the sample norm for any of the hard tissue measurements of Tweed, Steiner, Wits, Jarabak, or Ricketts. A **strong negative correlation** was indicated between degree of soft tissue deviation from Ricketts' E-line for both upper and lower lips and individual case TAS for favorability, i.e. decreased deviation from the sample's norms were associated with increased favorability. **No significant correlations** of individual case TAS for favorability and degree of deviation from soft tissue measurements of Holdaway and Arnett were noted.

Ceph Analysis	Measurement	Original Norm	Sample Mean		Sample Range	
			T1	T2	T1	T2
Tweed	FMPA	25°	30.24	31.17	43.32—19.08	42.96—16.50
	IMPA	90°	95.13	95.98	112.95—77.15	124.26—77.80
	FMIA	65°	54.63	52.85	73.66—33.28	68.78—32.59
Steiner	SNA	82°	79.41	79.31	87.85—67.39	89.99—65.12
	SNB	80°	76.83	76.7	85.23—67.74	85.04—64.50
	ANB	2°	2.67	2.62	7.50–(-)2.42	8.27–(-)0.73
	U1 to NA	4mm	6.21	5.32	14.01–(-)2.78	12.51—0.07
	U1 to NA	22°	26.07	25.29	41.07–(-)0.92	41.63—7.37
	L1 to NB	4mm	6.32	6.29	15.27–(-)0.18	12.83—1.43
	L1 to NB	25°	28.86	29.39	48.93—19.43	47.85—12.15
	Interincisal angle	131°	122.42	122.7	145.69—95.12	157.18—88.44
	GoGn-SN	32°	36.97	36.72	47.60—24.86	48.03—23.52
Wits	Wits	-0.5mm	-0.32	0.09	6.95–(-)12.19	8.58–(-)10.01
Jarabak	Saddle angle (N-S-Ar)	123°	125.97	126.1	134.55—115.60	136.32—116.34
	Articular angle (S-Ar-Go)	143°	140.65	141.05	160.09—129.24	151.80—127.89
	Upper Gonial angle (Ar-Go-N)	54°	52.55	51.29	62.08—45.53	57.69—44.79
	Lower Gonial angle (N-Go-Me)	73°	77.8	78.28	88.92—64.09	90.49—62.70
	ACB/Corpus length ratio	100%	107%	106%	132—88	125—87
	PFH/AFH ratio	65%	64%	65%	71—58	72—57
	PCB/Ramus height ratio	75%	70%	67%	86—48	85—48
	Jarabak Axis	93°	95.89	96.96	87.26—106.92	89.31—103.76
Ricketts	Upper incisor axis (to Ba-N)	89°	89.74	90.88	117.37—69.50	108.19—75.00
	Facial plane (FH-N-Pog)	87°	83.96	83.28	95.10—75.48	93.04—73.01
	Facial convexity (A pt to N-Pog)	3mm	2.31	1.73	7.35–(-)2.68	9.00–(-)3.84
	L1 to A-Po	1mm	8.26	6.83	16.89—1.89	12.40—0.99
	Cranial deflection (FH - Ba-N)	27°	21.95	21.03	27.55—11.12	25.91—14.54
	Lower 1 to MPA	93°	95.13	95.98	112.95—77.15	124.26—77.80
	Facial taper (N-Po-Go)	68°	72.59	72.45	82.13—64.79	82.04—63.59
	E-line (Upper lip rel: to E-line)	-1mm	-1.82	-3.30	7.52–(-)12.17	4.58–(-)12.00
	E-Line (Lower lip ant to E-line)	1mm	-0.68	-2.39	8.16–(-)9.86	5.94–(-)13.44
Holdaway	H Line (lower lip)	0.5mm	0.9	0.35	6.37–(-)2.33	4.04—1.81
	H Angle (10° at 0mm convexity)	10°	15.8	14.84	3.05—28.85	3.92—25.85

Table 3: Cephalometric analyses, norms, and sample means and ranges

Ceph Analysis	Measurement	Original Norm	Sample Mean T1	Sample Mean T2	Sample Range T1	Sample Range T2
Arnett		**F/M**				
Projection to TVL (*AKA* FH perp to Sn):						
	Glabella' (Sglabella)	-8.5/-8.0	-1.48	-0.80	7.4--(-)10.92	10..96--(-)10.43
	Nasal projection (Mnosetip)	16.0/17.4	14.09	15.82	19.49---7.89	21.71---11.13
	Subnasale (Subnose)	0/0	0	0	0---0	0---0
	A' soft tissue (Usulcus)	-0.1/-0.3	-0.50	-1.18	5.73--(-)6.43	4.88--(-)5.43
	Upper lip ant (Ulipant)	3.7/3.3	0.57	-0.56	9.68--(-)8.88	8.85--(-)8.04
	Mx1 (U1 In)	-9.2/-12.1	-11.27	-13.72	(-)0.29--(-)22.28	(-)1.73--(-)24.11
	Md1 (L1 In)	-12.4/-15.4	-14.69	-16.45	(-)2.88--(-)23.72	(-)3.49--(-)26.64
	Lower lip anterior (Llipant)	1.9/1.0	-1.86	-3.15	9.35--(-)10.90	8.21--(-)13.15
	B' soft tissue (Lsulcus)	-5.3/-7.1	-10.82	-12.66	1.44--(-)21.16	(-)1.58--(-)25.58
	Pogonion, (Spogonion)	-2.6/-3.5	-10.44	-11.72	2.90--(-)22.80	(-)1.25--(-)24.64
Harmony values rel to TVL (*AKA* FH perp to Sn)						
Intra Md:	Md1-Pog' (L1 In-Spogonion)	9.8/11.9	4.25	4.88	15.58--(-)4.28	12.70--(-)3.05
	Lower lip ant to Pog' (Llipant-Spogonion)	4.5/4.4	-8.58	-8.50	(-)1.95--(-)15.42	(-)0.53--(-)15.33
	B' to Pog' (Lsulcus-Spogonion)	2.7/3.6	0.38	0.94	4.88--(-)3.34	5.39--(-)4.07
Interjaw:	Subnasale-Pog' (Subnose-Spogonion)	3.2/4.0	10.44	11.72	22.80--(-)2.90	24.64---1.25
	A'- B' (Usulcus-Lsulcus)	5.2/6.8	10.32	11.45	21.85--(-)2.13	20.72---2.16
	Upper lip ant-Lower lip ant (Ulipant-Llipant)	1.8/2.3	2.43	2.59	12.19--(-)3.80	6.38--(-)0.49
Total face:	Facial angle (Sglab-Subnose-Spog)	169.3/169.4	167.71	167.38	159.30---178.61	156.03---178.59
	Glabella'-A' (Sglabella-Usulcus)	8.4/7.8	0.98	-0.39	11.73--(-)12.47	10.57--(-)13.81
	Glabella'-Pog' (Sglabella-Spogonion)	5.9/4.6	-8.96	-11.05	12.45--(-)27.10	8.15--(-)32.30

Table 3 continued: Cephalometric analyses, norms, and sample means and ranges

	Task 1 Difficulty	Task 2 Favorability
Case #	Total Assessment Score	Total Assessment Score
101	9	8
102	9	10
103	18	11
104	11	18
105	15	10
106	14	7
107	9	8
108	9	11
109	13	18
110	11	17
111	8	18
112	18	8
201	6	17
202	13	12
203	15	16
204	12	6
205	13	7
206	16	6
207	7	17
208	12	10
209	14	11
210	10	10
211	12	16
212	14	16
301	7	11
302	12	13
303	15	8
304	8	12
305	10	8
306	15	18
307	14	9
308	16	16
309	9	13
310	17	18
311	14	11
312	7	7

Table 4: Total Assessment Score (TAS): Total of 6 judges' subjective rank scores of 1, 2, or 3 for each case (total scale of 18). Higher TAS equals increased T1 difficulty and increased T2 favorability. Yellow highlighted cases are those with unanimous assessments of greatest difficulty at T1, and greatest favorability at T2.

Measurement	Published Norm	T1 Difficulty		T2 Favorability	
		r	p	r	p
Tweed					
FMPA	25°	0.09	0.63	-0.04	0.79
IMPA	90°	-0.12	0.49	-0.07	0.69
FMIA	65°	0.02	0.91	0.12	0.49
Steiner					
SNA	82°	0.13	0.44	-0.10	0.57
SNB	80°	0.18	0.33	-0.05	0.77
ANB	2°	0.23	0.19	0.00	0.99
U1-NA mm	4mm	-0.04	0.84	-0.16	0.36
U1-NA degrees	22°	-0.13	0.47	-0.21	0.24
L1-NB mm	4mm	0.01	0.94	0.06	0.75
L1-NB degrees	25°	-0.15	0.39	-0.02	0.91
Interincisal angle	131°	-0.16	0.38	-0.08	0.66
GoGn-Sn	32°	0.26	0.14	0.01	0.97
Wits	-0.5mm	0.14	0.43	-0.19	0.28
Jarabak					
Saddle angle	123°	-0.08	0.64	-0.06	0.73
Articular angle	143°	-0.18	0.92	0.03	0.87
Upper Gonial angle	54°	0.19	0.29	-0.03	0.88
Lower Gonial angle	73°	0.18	0.30	0.07	0.71
ACB/Corpus length ratio	100%	0.24	0.17	0.17	0.34
PFH/AFH ratio	65%	-0.01	0.95	-0.12	0.49
PCB/Ramus height ratio	75%	0.06	0.72	0.23	0.20
Jarabak Axis	93°	0.02	0.92	-0.04	0.83
Ricketts					
Upper incisor axis	89°	-0.10	0.55	0.10	0.59
Facial plane	87°	0.05	0.77	0.12	0.48
Facial convexity	3mm	0.20	0.26	-0.30	0.09
L1 to A-Pog	1mm	-0.04	0.84	0.07	0.70
Cranial deflection	27°	-0.08	0.65	-0.14	0.43
L1 to MPA	93°	-0.12	0.49	-0.07	0.69
Facial Taper	68°	-0.15	0.41	-0.20	0.26
E-line (upper lip to E-line	-1mm	0.16	0.36	-0.30	0.08
E-line (lower lip to E-line	1mm	0.20	0.26	-0.38	0.03
Holdaway					
H Line (lower lip)	0.5mm	0.15	0.39	-0.35	0.04
H Angle	10°	0.02	0.90	-0.16	0.38

Table 5: Correlations of Total Assessment Score with Published Norms Yellow highlighted numbers show significant correlations at p<0.05. Numbers noted in green show a tendency toward significant correlation, but fall above the 0.05 p-value threshold for significance.

Measurement	Published Norm	T1 Difficulty		T2 Favorability	
		r	p	r	p
Arnett	**F/M**				
Projection to TVL (*AKA* FH perp to SN)					
Glabella'	-8.5/-8.0	-0.02	0.90	-0.03	0.86
Nasal Projection	16.0/17.4	-0.07	0.69	0.16	0.37
Subnasale	0/0	n/a	n/a	n/a	n/a
A' soft tissue	3.7/3.3	0.02	0.91	-0.22	0.22
Upper lip anterior	-0.1/-0.3	0.07	0.70	-0.14	0.43
Mx1	-9.2/-12.1	-0.06	0.72	0.00	0.99
Md1	-12.4/-15.4	0.20	0.24	-0.11	0.52
Lower lip anterior	-5.3/-7.1	0.30	0.08	-0.14	0.44
B' soft tissue	1.9/1.0	0.26	0.13	-0.09	0.60
Pogonion'	-2.6/-3.5	0.26	0.13	-0.05	0.78
Harmony values to TVL (*AKA* FH perp to SN)					
Intra Md:					
Md1-Pog'	9.8/11.9	0.08	0.65	0.09	0.63
Lower lip ant to Pog'	4.5/4.4	0.06	0.74	0.09	0.60
B' to Pog'	2.7/3.6	0.16	0.37	0.12	0.51
Interjaw:					
Subnasale-Pog'	3.2/4.0	0.26	0.14	-0.04	0.84
A'-B'	5.2/6.8	0.23	0.19	-0.11	0.56
Upper lip ant-Lower lip ant	1.8/2.3	0.30	0.09	-0.20	0.26
Total Face:					
Facial angle	169.3/169.4	0.30	0.09	-0.04	0.81
Glabella'-A'	8.4/7.8	0.00	0.99	-0.08	0.65
Glabella'-Pog'	5.9/4.6	0.18	0.31	-0.05	0.76

Table 5 Cont: Correlations of Total Assessment Score with Published Norms

Measurement	Sample Mean		T1 Difficulty		T2 Favorability	
	T1	T2	r	p	r	p
Tweed						
FMPA	30.24	31.17	-0.39	0.02	0.01	0.95
IMPA	95.13	95.98	-0.03	0.85	-0.11	0.54
FMIA	54.63	52.85	0.21	0.22	0.06	0.72
Steiner						
SNA	79.41	79.31	0.18	0.29	-0.13	0.48
SNB	76.83	76.7	0.00	0.99	0.02	0.91
ANB	2.67	2.62	0.22	0.21	-0.04	0.81
U1-NA mm	6.21	5.32	0.14	0.44	-0.04	0.82
U1-NA degrees	26.07	25.29	-0.07	0.71	-0.24	0.18
L1-NB mm	6.32	6.29	0.15	0.42	-0.02	0.93
L1-NB degrees	28.86	29.39	0.05	0.78	-0.20	0.25
Interincisal angle	122.42	122.7	0.05	0.76	-0.16	0.35
GoGn-Sn	36.97	36.72	-0.04	0.81	-0.08	0.67
Wits	-0.32	0.09	0.14	0.43	-0.05	0.77
Jarabak						
Saddle angle	125.97	126.1	-0.06	0.73	-0.01	0.95
Articular angle	140.65	141.05	-0.06	0.74	0.09	0.61
Upper Gonial angle	52.55	51.29	0.04	0.84	-0.29	0.10
Lower Gonial angle	77.8	78.28	-0.07	0.71	-0.03	0.89
ACB/Corpus length ratio	107%	106%	0.08	0.66	0.09	0.60
PFH/AFH ratio	64%	65%	-0.07	0.68	-0.12	0.49
PCB/Ramus height ratio	70%	67%	0.25	0.16	-0.12	0.49
Jarabak Axis	95.89	96.96	0.59	0.74	-0.12	0.51
Ricketts						
Upper incisor axis	89.74	90.88	-0.12	0.49	0.10	0.56
Facial plane	83.96	83.28	-0.15	0.39	0.19	0.27
Facial convexity	2.31	1.73	0.30	0.09	-0.17	0.34
L1 to A-Pog	8.26	6.83	0.15	0.39	-0.07	0.69
Cranial deflection	21.95	21.03	0.11	0.53	0.20	0.25
L1 to MPA	95.13	95.98	-0.03	0.86	-0.11	0.54
Facial Taper	72.59	72.45	0.00	0.99	-0.08	0.63
E-line (upper lip to E-line)	-1.82	-3.30	0.16	0.37	-0.45	0.008
E-line (lower lip to E-line)	-0.68	-2.39	0.10	0.59	-0.51	0.002
Holdaway						
H Line (lower lip)	0.9	0.35	0.23	0.19	-0.32	0.07
H Angle	16.22	14.96	0.20	0.26	-0.31	0.08

Table 6: Correlations of Total Assessment Score with Sample means Yellow highlighted numbers show significant correlations at $p<0.05$. Numbers noted in green show a tendency toward significant correlation, but fall above the 0.05 p-value threshold for significance.

Measurement	Sample Mean T1	Sample Mean T2	T1 Difficulty r	T1 Difficulty p	T2 Favorability r	T2 Favorability p
Arnett						
Projection to TVL (*AKA* FH perp to SN)						
Glabella'	-1.48	-0.80	-0.11	0.54	0.33	0.06
Nasal Projection	14.09	15.82	-0.14	0.43	0.12	0.50
Subnasale	0	0	n/a	n/a	n/a	n/a
A' soft tissue	-0.50	-1.18	-0.02	0.89	-0.19	0.27
Upper lip anterior	0.57	-0.56	-0.04	0.82	-0.18	0.32
Mx1	-11.27	-13.72	-0.16	0.35	0.03	0.85
Md1	-14.69	-16.45	0.05	0.78	-0.07	0.68
Lower lip anterior	-1.86	-3.15	0.11	0.55	-0.08	0.65
B' soft tissue	-10.82	-12.66	0.04	0.80	0.04	0.82
Pogonion'	-10.44	-11.72	0.00	0.99	0.07	0.69
Harmony values to TVL (*AKA* FH perp to SN)						
Intra Md:						
Md1-Pog'	4.25	4.88	0.08	0.66	-0.01	0.97
Lower lip ant to Pog'	-8.58	-8.50	0.12	0.49	-0.21	0.23
B' to Pog'	0.38	0.94	-0.06	0.73	-0.10	0.57
Interjaw:						
Subnasale-Pog'	10.44	11.72	0.00	0.99	0.07	0.69
A'-B'	10.32	11.45	0.15	0.39	-0.04	0.84
Upper lip ant-Lower lip an	2.43	2.59	0.27	0.12	-0.11	0.55
Total Face:						
Facial angle	167.71	167.38	0.21	0.23	-0.12	0.52
Glabella'-A'	0.98	-0.39	-0.20	0.24	0.31	0.07
Glabella'-Pog'	-8.96	-11.05	-0.14	0.44	0.26	0.15

Table 6 Cont.: Correlations of Total Assessment Score with Sample means

DISCUSSION

The main objective of this study was to determine the level of agreement between published cephalometric norms and clinician's subjective assessments. That is to say, the objective was to assess the utility of these norms as an objective means of validating subjective views. The following points were demonstrated by our results:

1. The results did not support our hypotheses that decreased deviations from hard tissue norms would be correlated with decreased difficulty and increased favorability. As can be seen by the lack of significant correlations, the judges' views of treatment difficulty and treatment success were not highly associated with patients' coincidence with hard tissue cephalometric norms. One measure, FMPA, paradoxically appeared to show that greater deviation from the norm is associated with assessments of less difficult treatment!
2. The results demonstrated fair support for our hypothesis that decreased soft tissue deviations are correlated with increased favorability (no support for decreased difficulty). Some traditional soft tissue norms, E-line and H-line, appeared to reflect, to a limited degree (0.51 at greatest correlation), the judges' views of treatment success.

If we assume that our norms are valid, why wasn't there greater agreement and support for our hypotheses?

The first way to address this question is to look at the limitations inherent in this study to see where errors introduced may have affected the correlations. These limitations included: 1. demographics of the sample, 2. error and variation in measurement, 3. error associated with acquiring the subjective data, and 4. limitations in the type of data acquired.

A limitation of the sample was the inclusion of patients of varied age groups, sexes, and ethnicities. Cephalometric norms vary greatly for different ages, sexes, and ethnicities. Our sample's mean age was 17.30, it included fewer boys than girls, and its ethnicity was representative of a large, multicultural, urban center. Therefore, comparing our sample's deviations to published norms that were derived from samples of Caucasian adolescent boys, may have questionable validity.

This study was also limited by the technique of measuring deviations from norms. First, the study required the use of single, discrete numbers for all of our cephalometric norms. Many of the published norms that were used were given as a range, with modifying factors based on age or skeletal variations (e.g. Jarabak and Holdaway). In order to reasonably limit our correlation analysis, it was necessary to choose the median number in these ranges as the definitive norm for measuring deviations. This introduced some error by resolving a norm into a single number, which may never have been intended as universal. Secondly, some approximations were required when measuring deviations from specific norms. For example, Arnett's soft tissue measurements should be made based on a constructed "True Vertical Line" on headfilms taken in natural head position. The headfilms used could not be oriented in this fashion and an approximation was used (perpendicular line to Frankfort-Horizontal through subnasale). This may have added an unknown quantity of variability to our measurements of deviations.

Acquiring subjective data for our sample was a difficult task. The grouping of our sample into 3 groups of 12 cases, though necessary for logistics, may have introduced some limitations. Requiring the judges to rank the 12 cases in each group independently of the other groups probably introduced artificial ranking assessments of difficulty and favorability. As evidence, one judge noted during the study that "all the cases in that [second] group are difficult". Artificial "forced" rankings may have skewed the TAS and could have led to less significant correlations. Assessing and ranking all 36 cases as one group may have avoided this limitation and may or may not have resulted in more significant correlations.

The most significant uncertainty of this study is probably presented by the premise that subjective assessments can be distilled down to a single, accurate, representative rank score for use in correlations. While individual cephalometric deviations can be accurately determined, subjective assessments of casts and photos defy simple measurement and comparison. We were limited in our ability to determine which aspects of the case records the judges relied on for their decisions or what their criteria for difficulty or favorability were. The inclusion of a comment area of the Subjective Assessment Sheets provided some insight into this limitation. It was clear that some degree of variability was present, and there was more variability in assessment of difficulty, rather than favorability. This was shown by the limited unanimous assessments for difficulty (TAS=18, only 2 cases). The judges' written comments on the Subjective Assessment Sheets for assessing favorability also show some variability: e.g. "good smile", "midlines-on", "good profile", "good overbite/overjet", "good interdigitation", and "good class 1 canines". However, the fact that the judges' resultant assessments for favorability were not more disparate, as evidenced by the high number (5 cases) of unanimously assessed "most favorable cases", suggests that our judges/clinicians may be looking at records in similar ways and making consistently similar decisions. Previous published studies of correlations between subjective and objective assessments provided minimal information and insight regarding what clinicians were actually looking at during assessments. This study is somewhat unique in having used full records and documented subjective assessments (via additional comments) in a detailed manner. Although difficult to quantify, subjective assessments (especially for favorability) gathered and ranked in the manner used in this study, appear to have a degree of consistency that makes them valid for correlations.

While there were several limitations that may have led to the lack of correlations, another logical reason for the lack of agreement may simply be that published norms are not valid for determining specific difficulty or favorability. Evidence for this rational is provided by previous studies of correlations of assessments with norms. Wylie's study showed very low correlation between deviations in FMIA (relative to Tweed's norm) and esthetic straightening of the profile (angle of convexity), and low correlation between maxillary incisor position changes and profile straightening.[19] Park and Burstone also found low correlations between objective facial esthetics (accepted soft tissue norms) and coincidence with a norm for lower incisor position to A-Po.[23] Similarly, DuClos found poor correlations of clinician' subjective assessments of positive facial esthetics with degree of deviations from hard tissue norms.[24] In summary, these studies, similar to this study, showed modest to no correlations between coincidence with published norms and improved facial esthetics. Clearly, the fact that treatment planning using published norms as a goal has been greatly discarded in orthodontics is recognition of inherent limitations in the use of norms.

The greatest value of this study is the evidence that there were no highly significant correlations of norms with subjective assessments. It is important to understand the limitations of using published norms. Although correctly used for description and comparison of malocclusions and interclinician communication, use of traditional published norms cannot provide guarantees of easy treatment plans or more successful treatments. As E.H. Hixon wrote in 1956, "It is abusing the norm to use it alone for evaluation in diagnosis, or to use the average as an objective in treatment planning. A norm is not a substitute for professional judgement."[33] Norms are excellent diagnostic aids for description of existing problems, but individual patient variation requires us to make subjective judgements for proper treatment.

BIBLIOGRAPHY

1. Broadbent BH. A new x-ray technique and its application to orthodontia. Angle Orthod. 1931;1:45-66.
2. Tweed C. Indications for the extraction of teeth in orthodontic procedure. Am J Orthod. 1944;30:401-460.
3. Tweed C. The Frankfort-mandibular plane angle in orthodontic diagnosis, classification, treatment planning, and prognosis. Am J Orthod and Oral Surg. 1946;32:175-230.
4. Downs WB. Variations in facial relationships: Their significance in treatment planning and prognosis. Am J Orthod and Oral Surg. 1948;34:812-840.
5. Bjork A. The face in profile. Svensk Tandlakare-Tiddskrift. 1947;40.
6. Steiner C. Cephalometrics for you and me. Am J Orthod. 1953;39:729-755.
7. Sassouni V. A roentographic cephalometric analysis of cephalo-facial-dental relationships. Am J Orthod. 1955;41:735-764.
8. Jacobson A, The "Wits" appraisal of jaw disharmony. Am J Orthod. 1975;67:125-138.
9. Burstone CJ. The Integumental Profile. Am J Orthod. 1958;44:1-25.
10. Ricketts RM. Cephalometric synthesis. Am J Orthod. 1960;46:647-673.
11. Ricketts RM. Perspectives in the clinical application of cephalometrics – The first fifty years. Angle Orthod. 1981;51:115-150.
12. Merrifield LL. The profile line as an aid in critically evaluating facial esthetics. Am J Orthod. 1966;52:804-822.
13. Holdaway RA. A soft tissue cephalometric analysis and its use in orthodontic treatment planning. Part I. Am J Orthod. 1983;84:1-28.

14. Holdaway RA. A soft tissue cephalometric analysis and its use in orthodontic treatment planning. Part II. Am J Orthod. 1984;85:279-293.

15. Arnett GW. Facial keys to orthodontic diagnosis and treatment planning. Part I. Am J Orthod Dentofac Orthop. 1993;103:299-312.

16. Arnett GW. Facial keys to orthodontic diagnosis and treatment planning. Part II. Am J Orthod Dentofac Orthop. 1993;103:395-411.

17. Arnett GW. Soft tissue cephalometric analysis: Diagnosis and treatment planning of dentofacial deformity. Am J Orthod Dentofac Orthop. 1999;116:239-253.

18. Riedel RA. Esthetics and its relation to orthodontic therapy. Angle Orthod. 1950;20:168-178.

19. Wylie LW. The mandibular incisor – Its role in facial esthetics. Angle Orthod. 1955;25:32-41.

20. Reidel RA. An analysis of dentofacial relationships. Am J Orthod. 1957;43:103-119.

21. Poulton DR. Facial esthetics and angles. Angle Orthod. 1957;27:133-137.

22. Peck H. A concept of facial esthetics. Angle Orthod. 1970;40:284-318.

23. Park YC. Soft-tissue profile – Fallacies of hard-tissue standards in treatment planning. Am J Orthod Dentofac Orthop. 1986;90:52-62.

24. DuClos JB. Factors in clinical decision making: Associations between subjective evaluations by clinicians and objective evaluations by physical measurement. Unpublished thesis. 1995.

25. Teeter D. The extraction decision. A pilot study using clinician preferences in a retrospective analysis. Unpublished thesis. 1997.

26. Baumrind S. Computer aided head film analysis: The University of California, San Francisco method. Am J Orthod. 1980;78:41-80.

27. Nalchajian N. Facial esthetics in borderline extraction and non-extraction patients. Associations between objective cephalometric measurements and subjective assessments of expert clinicians. Unpublished thesis. 1999.

28. Baumrind S. The reliability of head film measurements. 1. Landmark Identification. Am J Orthod. 1971;60:111-127.

29. Dawson-Saunders B. Basic and clinical biostatistics. Newark: Appleton & Lange. 1994;54.

30. Han UK. Consistency of orthodontic treatment decisions relative to diagnostic records. Am J Orthod Dentofac Orthop. 1991;100:212-219.

31. Silling G. The significance of cephalometrics in treatment planning. Angle Orthod. 1979;49:259-263.

32. Foster EJ. Profile preferences among diversified groups. Angle Orthod. 1973;43:34-40.

33. Hixon EH. The norm concept and cephalometrics. Am J Orthod. 1956;42:898-906.

34. Hixon EH. Cephalometrics: A perspective. Angle Orthod. 1972;42:200-209.